OIPE IAP88 NOV 14 2005 PATENT & TRADEMARK OFFICE

AF/2655 JFW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor : David N. Weise	Appeal No.
Appln. No.: 09/903,055	
Filed : July 11, 2001	Group Art Unit: 2655
For : METHOD AND APPARATUS FOR PARSING TEXT USING MUTUAL INFORMATION	Examiner: Rivero, M.
Docket No.: M61.12-0349	

TRANSMITTAL OF APPEAL BRIEF
(PATENT APPLICATION - 37 C.F.R. §41.37)

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450, THIS 9th DAY OF November, 2005.

PATENT ATTORNEY

Sir:

Transmitted herewith is the Appeal Brief in this application with respect to the Notice of Appeal filed on October 25, 2005.

FEE STATUS

[---] Small entity status under 37 C.F.R. §§ 1.9 and 1.27 is established by a verified statement.

FEE FOR FILING APPEAL BRIEF

Pursuant to 37 C.F.R. §41.20(b)(2) the fee for filing the Appeal Brief is $500.00.

The Director is authorized to charge any additional fees associated with this paper or credit any overpayment to Deposit Account No. 23-1123. A duplicate copy of this communication is enclosed.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: ______________________
Theodore M. Magee, Reg. No. 39,758
Suite 1400 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

TMM:sew

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor : David N. Weise		Appeal No.
Appln. No. : 09/903,055		
Filed : July 11, 2001		Group Art Unit: 2655
For : METHOD AND APPARATUS FOR PARSING TEXT USING MUTUAL INFORMATION		Examiner: Rivero, M.
Docket No.: M61.12-0349		

BRIEF FOR APPELLANT

Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450, THIS 9th DAY OF November, 2005.

PATENT ATTORNEY

Sir:

This is an appeal of an Office Action dated July 25, 2005 finally rejecting claims 1-3, 5-8, 10, 12-19, 21 and 22.

REAL PARTY IN INTEREST

Microsoft Corporation, a corporation organized under the laws of the state of Washington, and having offices at One Microsoft Way, Redmond, Washington 98052, has acquired the entire right, title and interest in and to the invention, the application, and any and all patents to be obtained therefor, as set forth in the Assignment filed with the parent application of this patent application and recorded on Reel 11991, frame 0793.

RELATED APPEALS AND INTERFERENCES

There are no known related appeals or interferences that will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

STATUS OF THE CLAIMS

I.	Total number of claims in the application.		
		Claims in the application are:	1-3, 5-8, 10, 12-19, 21, and 22
II.	Status of all the claims.		
	A.	Claims cancelled:	4, 9, 11, 20
	B.	Claims withdrawn but not cancelled:	---
	C.	Claims pending:	1-3, 5-8, 10, 12-19, 21, and 22
	D.	Claims allowed:	---
	E.	Claims rejected:	1-3, 5-8, 10, 12-19, 21, and 22
	F.	Claims Objected to:	21
III.	Claims on appeal		
		The claims on appeal are:	1-3, 5-8, 10, 12-19, 21, and 22

STATUS OF AMENDMENTS

An Amendment After Final was filed September 13, 2005 to change the dependency of claim 21 from canceled claim 20 to claim 19. The Advisory Action of October 4, 2005 did not indicate whether the amendment would or would not be entered into the record. Appellants have assumed the amendment was entered.

SUMMARY OF INVENTION

Independent claim 1 provides a method of generating 312 a score for a node identified during a parse (Fig. 3) of a text segment 302. Generating the score involves identifying a phrase level (page 16, line 9 - page 18, line 27), identifying a word

class (page 14, line 30 - page 16, line 6) for at least one word that neighbors text spanned by the node, and generating a score by determining a mutual information metric (page 18, line 29 - page 20, line 21) based on the phrase level and the word class.

Independent claim 10 provides a parser (Fig. 3, page 12, line 21 - page 14, line 27) for generating a syntax structure from a text segment 302. The parser includes a seeding unit 305 for inserting words from the text segment into a candidate list 306. A node selector 307 promotes nodes from the candidate list 306 to a node chart 308. A rule engine 309 combines nodes in the node chart 308 to form a larger node. The larger node is scored by a metric calculator 312, where the score is based in part on mutual information (page 18, line 29 - page 20, line 21) determined based on a phrase level (page 16, line 9 - page 18, line 27) of the node and at least one word in the text segment.

Independent claim 19 provides a computer-readable medium having computer-executable instructions for performing steps that include dividing 303 a text segment 302 into words and forming 309 syntax nodes that each represent a syntax structure for one or more words. A syntax node is scored 312 to indicate its likelihood of appearing in a full parse structure for the text segment where the score is a mutual information score (page 18, line 29 - page 20, line 21) that is based in part on a phrase level (page 16, line 9 - page 18, line 27) of the syntax node. The score for the syntax node is then used 307 when forming the full parse structure.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-3 and 6-8 are rejected under 35 U.S.C. §102(b) as being anticipated by Su et al. (U.S. Patent No. 5,418,717, hereinafter Su).

Claims 5, 10, 12-19, 21 and 22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera et al. (U.S. Patent No. 4,868,750, hereinafter Kucera).

ARGUMENT

Claims 1-3, 6 and 8

Claims 1-3, 6 and 8 were rejected under 35 U.S.C. §102(b) as being anticipated by Su.

Independent claim 1 provides a method of generating a score for a node identified during a parse of a text segment. Generating the score involves identifying a phrase level, identifying a word class for at least one word that neighbors the text spanned by the node, and generating a score by determining a mutual information metric based on the phrase level and the word class.

Su does not show or suggest the invention of claim 1 because it does not show or suggest determining a mutual information metric based on a phrase level and a word class.

In the Final Office Action, it was asserted that Su suggests employing a mutual information metric as part of the parsing process in that it uses a conditional probability for determining its syntactic score. (Citing col. 5, lines 8-14, col. 9, lines 32-40 and Col. 13, lines 34-37). However, Su never refers to this syntactic score as mutual information and, as is well known in the art, a conditional probability is not the same as mutual information.

In particular, a conditional probability does not provide a correlation between two or more events like mutual information does. A conditional probability simply provides the probability of a first event given a second event. This is not a correlation because it does not indicate how the first event relates to the second event or visa versa.

To understand the difference, which is well known in the art, imagine a class of 20 children: 7 with red hair and freckles, 3 with red hair and no freckles, 7 with blond hair and freckles, and 3 with blond hair and no freckles. In this class,

the conditional probability of seeing a child with freckles given that you see a child with red hair is 0.7 (Number of children with freckles and red hair/number of children with red hair). Although this probability indicates that it is more likely than not that a child will have freckles if they have red hair, red hair is not correlated to freckles. The reason for this is that the probability of a child having freckles regardless of their hair color is also 0.7. (Number of children with freckles/number of children). Thus, for this example, we cannot say that red hair and freckles are related to each other in any way. In fact, in this example, red hair and freckles are completely independent of each other because a child with freckles is just as likely to have blond hair as to have red hair. Thus, there is no correlation between red hair and freckles in the example above.

In the example above, a mutual information metric would be zero. In particular, it would be computed as:

$$I = \log \frac{p(redhair\ freckles)}{p(redhair)p(freckles)} = \log \frac{(7/20)}{(10/20)(14/20)} = \log 1 = 0$$

where I is the mutual information, $p(redhair\ freckles)$ is the probability of red hair and freckles occurring together, $p(redhair)$ is the probability of red hair, and $p(freckles)$ is the probability of freckles. As is well known in the art, a value of zero for mutual information indicates that the variables are independent of each other meaning there is zero correlation between the two variables.

If we change the example above so that the population of children in the class of 20 is: 7 with red hair and freckles, 3 with red hair and no freckles, and 10 with blond hair and no freckles, the conditional probability of freckles given red hair remains 0.7. However, it can be seen that in this example, red hair provides more information about whether a child has freckles. In particular, if we know that a child has red hair,

we know it is more likely that the child will have freckles than if they are blond. Thus, in this example, there appears to be a relationship or correlation between red hair and freckles. However, the conditional probability is the same as when there was no relationship between red hair and freckles. Thus, the conditional probability does not indicate this correlation.

Mutual information indicates the relationship between red hair and freckles in the second example by providing a different value than in the first example. In particular, the mutual information metric in the second example becomes:

$$I = \log \frac{p(redhair\ freckles)}{p(redhair)p(freckles)} = \log \frac{(7/20)}{(10/20)(7/20)} = \log 2 = 0.301$$

which indicates a positive relationship between red hair and freckles.

As the examples above show, conditional probability is not the same as mutual information because conditional probability does not show the correlation between two or more events.

Because Su does not mention mutual information and because conditional probability is not the same as mutual information, Su does not anticipate claim 1 or claims 2, 3, 6, and 8, which depend from claim 1.

Claim 5

Claim 5 was rejected under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera. Claim 5 depends from claim 1.

As noted above, Su does not show or suggest determining a mutual information metric as found in claim 1. Similarly, Kucera does not mention determining a mutual information metric.

The Final Office Action asserted that Kucera shows a mutual information metric at column 2, lines 20-34 when it discusses its collocation probability. However, the collocation

probability in Kucera is not a mutual information metric because it does not indicate the correlation between two events. This can be seen using a simple set of training data. If in the set of training data, two events occurred next to each other ten times and the individual events each occurred ten times, the collocation function in Kucera (see column 12, lines 30-40) would evaluate to 10/(10x10) or 1/10. If we doubled the size of the training data simply by duplicating the set of training data, the score would then become 20/(20x20) or 1/20. However, since the same training data is found in both examples, the correlation between the two events should be the same in both sets of training data. The fact that the collocation function in Kucera changes in this example, indicates that it does not show the correlation between two events but instead is providing some other type of information. As such, the function shown in Kucera is not a mutual information metric.

In the Final Office Action, the Examiner pointed to the background of Lu (U.S. Patent No. 5,819,260), which makes a statement that "[t]he statistics used include collocation information or mutual information, i.e., the probability that a given pair of a part-of-speech tags or a given pair of words tends to appear together in a given data collection." Lu goes on to cite Kucera directly later in the same paragraph. However, Lu is not saying that Kucera uses mutual information. Further, Lu is not saying that collocation information is the same as mutual information. This is particularly true in Kucera where the collocation determination made by Kucera is clearly not a mutual information metric.

Thus, in the combination of cited references, including Lu, there is no suggestion for generating a mutual information metric based on a phrase level. Su discusses conditional probabilities and Kucera discusses collocation scores, neither of which are mutual information metrics. Further, none of these

references indicate that mutual information could be used in either Su or Kucera instead of the conditional probabilities or the collocation scores.

Since none of the cited references show a mutual information metric generated based on a phrase level, and Su and Kucera do not even discuss mutual information, claim 5 is patentable over the combination of Su and Kucera.

Claim 7

Claim 7 was rejected under 35 U.S.C. §102(b) as being anticipated by Su.

Claim 7 depends indirectly from claim 1 and includes a further limitation wherein generating a score involves generating a score based on all possible word classes identified for a word to the left of the text spanned by the node and all possible word classes identified for a word to the right of the text spanned by the node.

In the Final Office Action, the Examiner cited column 17, lines 47-66; column 11, lines 46-50; and FIG. 7 of Su as showing this limitation. However, the cited sections do not show or suggest generating a score based on all possible word classes for a word. Instead, the cited sections show generating separate scores for each of a set of possible sequences of parts-of-speech. Each score is generated using a single part-of-speech for each word. The separate scores are never combined together. Thus, Su does not show a score that is generated based on all possible word classes for a word. Instead, it shows a plurality of scores that are each generated based on a single part-of-speech for each word.

This is substantially different from claim 7 where a score is generated based on all possible word classes identified for a word. As such, claim 7 is additionally patentable over Su.

Claims 10, 12, and 14-18

Claims 10, 12 and 14-18 were rejected under 35 U.S.C.

§103(a) as being unpatentable over Su in view of Kucera.

Independent claim 10 provides a parser for generating a syntax structure from a text segment. The parser includes a seeding unit for inserting words from the text segment into a candidate list. A node selector promotes nodes from the candidate list to a node chart. A rule engine combines nodes in the node chart to form a larger node. The larger node is scored by a metric calculator, where the score is based in part on mutual information determined based on a phrase level of the node and at least one word in the text segment.

The combination of Su and Kucera does not show or suggest the invention of claim 10 because neither Su nor Kucera mention mutual information and in particular, neither reference discusses mutual information that is determined based on a phrase level of a node.

Su uses conditional probabilities and Kucera uses collocation. As noted above, neither of these is mutual information. As such, the combination of Su and Kucera does not show or suggest mutual information that is determined based on a phrase level as found in claim 10. Claim 10 and claims 12 and 14-18, which depend therefrom, are therefore patentable over the combination of Su and Kucera.

Claim 13

Claim 13 was rejected under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera.

Claim 13 depends indirectly from claim 10 and includes a further limitation wherein the mutual information is determined based on all possible word classes for a word in the text segment.

In the Final Office Action, it was asserted that this limitation was shown in Kucera at Col. 1, line 65 - Col. 2, line 3 and in Su at Col. 17, lines 47-66; Col. 11, lines 46-50; and Figure 7.

The cited section of Kucera discusses identifying a

sequence of part-of-speech tags from a set of possible sequences by selecting a most probable sequence of part-of-speech tags. It does not mention mutual information or determining a value based on all possible word classes for a word in a text segment. Thus, Kucera does not show the limitation of claim 13.

Similarly, the cited sections of Su do not show this limitation. As noted above, the cited sections of Su generate a plurality of scores, with each score being based on a single part of speech for each word. These scores are never combined and as a result, Su does not generate a value based on all possible word classes for a word. In addition, since Su does not mention mutual information, it does not show or suggest determining mutual information based on all possible word classes for a word.

Since none of the cited sections of Su and Kucera shows or suggests determining mutual information based on all possible word classes for a word, the combination of Su and Kucera does not show or suggest the invention of claim 13.

Claim 19

Claim 19 was rejected under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera.

Independent claim 19 provides a computer-readable medium having computer-executable instructions for performing steps that include dividing a text segment into words and forming syntax nodes that each represents a syntax structure for one or more words. A syntax node is scored to indicate its likelihood of appearing in a full parse structure for the text segment where the score is a mutual information score that is based in part on a phrase level of the syntax node. The score for the syntax node is then used when forming the full parse structure.

As discussed above, neither Su nor Kucera show or suggested a mutual information score and in particular, neither reference shows or suggests a mutual information score based in

part on a phrase level of a syntax node. Su uses conditional probabilities and Kucera uses a collocation score, neither of which are mutual information.

Since neither Su nor Kucera show or suggest a mutual information score that is based on a phrase level of a node, their combination does not show or suggest the invention of claim 19.

Claims 21-22

Claims 21 and 22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera.

Claims 21 and 22 depend from claim 19 and include a further limitation wherein the mutual information score is based on all possible word classes of a word in the text segment.

As noted above for claim 13, neither Su nor Kucera shows a mutual information score that is based on all possible word classes for a word. As such, claims 21 and 22 are further patentable over the combination of Su and Kucera.

Conclusion

In light of the arguments above, Appellants request that the Board reverse the Examiner's rejection of claims 1-3, 5-8, 10, 12-19, 21 and 22.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: ______________________________

Theodore M. Magee, Reg. No. 39,758
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone:(612) 334-3222 Fax:(612) 334-3312

TMM:sew

Appendix A - Claims on Appeal

1. A method of generating a score for a node identified during a parse of a text segment, the method comprising:

identifying a phrase level for the node;

identifying a word class for at least one word that neighbors a text spanned by the node; and

generating a score by determining a mutual information metric based on the phrase level and the word class.

2. The method of claim 1 wherein identifying a word class comprises:

identifying a word class for a word to the left of the text spanned by the node; and

identifying a word class for a word to the right of the text spanned by the node.

3. The method of claim 2 wherein generating a score comprises generating a score based on the phrase level of the node, the word class of the word to the right of the text spanned by the node and the word class of the word to the left of the text spanned by the node.

5. The method of claim 2 wherein determining a mutual information metric comprises determining a mutual information metric based on the phrase level of the node, the word class of the word to the right of the text spanned by the node and the word class of the word to the left of the text spanned by the node.

6. The method of claim 2 wherein identifying a word class

further comprises:

identifying all possible word classes for a word to the left of the text spanned by the node; and

identifying all possible word classes for a word to the right of the text spanned by the node.

7. The method of claim 6 wherein generating a score comprises generating a score based in part on all of the identified word classes.

8. The method of claim 1 wherein identifying a word class comprises identifying all possible word classes for at least one word.

10. A parser for generating a syntax structure from a text segment, the parser comprising:

a seeding unit for inserting words from the text segment into a candidate list as nodes;

a node selector for promoting nodes from the candidate list to a node chart;

a rule engine for combining nodes in the node chart to form a larger node; and

a metric calculator for generating a score for a node formed by the rule engine, the score being based in part on mutual information determined based on a phrase level of the node formed by the rule engine and at least one word in the text segment.

12. The parser of claim 10 wherein the mutual information is determined based on a word class for a word in the text segment.

13. The parser of claim 12 wherein the mutual information

is determined based on all possible word classes for a word in the text segment.

14. The parser of claim 12 wherein the mutual information is determined based on a word class for a word to the left of a set of words spanned by the node formed by the rule engine.

15. The parser of claim 14 wherein the mutual information is determined based additionally on a word class for a word to the right of the set of words spanned by the node formed by the rule engine.

16. The parser of claim 10 further comprising a lexicon look-up for determining parts of speech for words in the text segment.

17. The parser of claim 16 wherein the seeding unit inserts a node for each part of speech of each word in the text segment.

18. The parser of claim 17 wherein the seeding unit further inserts nodes representing the beginning of the text segment and the ending of the text segment.

19. A computer-readable medium having computer-executable instructions for performing steps comprising:

dividing a text segment into words;

forming syntax nodes that each represent a syntax structure for one or more words;

scoring a syntax node to indicate its likelihood of appearing in a full parse structure for the text segment, the score being a mutual information score that is based in part on a phrase level of the syntax node; and

using the score for the syntax node when forming the full parse structure.

21. The computer-readable medium of claim 19 wherein the mutual information score is further based on all possible word classes of a word in the text segment.

22. The computer-readable medium of claim 21 wherein the mutual information score is based on a word class of a word that is next to a word that the syntax node represents.